DOCUMENT RESUME

ED 229 932 EC 151 931

AUTHOR	Covert, Robert W.; Suarez, Tanya M.
TITLE	Constructing Questionnaires for Early Childhood Special Education Programs. TADScript Number 7.
INSTITUTION	North Carolina Univ., Chapel Hill. Technical Assistance Development System.
SPONS AGENCY	Special Education Programs (ED/OSERS), Washington, DC. Handicapped Children's Early Education Program.
PUB DATE	Sep 82
CONTRACT	300-80-0752
NOTE	57p.; Portions of this TADScript were adapted from Guidelines and Criteria for Constructing Questionnaires by R. W. Covert.
PUB TYPE	Guides - Non-Classroom Use (055) -- Tests/Evaluation Instruments (160)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Attitude Measures; Biographical Inventories; *Disabilities; Early Childhood Education; *Material Development; *Parent Attitudes; Participant Characteristics; Participation; *Questionnaires; Rating Scales; Special Programs; Surveys; *Teacher Attitudes

ABSTRACT

Guidelines and criteria for constructing questionnaires to evaluate early childhood special education programs are presented. The information may be used to identify parent attitudes and staff attitudes and needs. Attention is directed to the focus, title, purpose statement, directions, demographic questions, set of items and responses, and structure and format. Determining the focus of the questionnaire before its construction involves identifying the following: the purpose of the questionnaire, the persons who will use the results, and the specific group that will answer the questionnaire. In addition to a title that is concise, understandable, and reflects the content of the instrument, an introductory statement is needed that specifies the purpose, how the data will be used, anonymity, and motivators for respondents to participate. Directions for completing the instrument need to be written at the language level of the respondents and should indicate the approximate amount of time required to complete the questionnaire. The demographic section should request information limited to those variables that will be used in the data analysis. To assist in the actual writing of questionnaires, 15 suggestions are presented for developing rating scales, along with tips on writing qualitative items (e.g., items which elicit information on job classification, sex); and supply items (i.e., the respondent produces the response rather than selecting an alternative provided on the questionnaire). A sample questionnaire is included. (SEW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED229932

CONSTRUCTING QUESTIONNAIRES FOR EARLY CHILDHOOD SPECIAL EDUCATION PROGRAMS*

Robert W. Covert
Evaluation Research Center
University of Virginia

Tanya M. Suarez
Technical Assistance Development System
University of North Carolina at Chapel Hill

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Kenneth Jani

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

September 1982

EC 151931

*Portions of this TADScript were adapted from Covert, R.W. Guidelines and Criteria for Constructing Questionnaires. Charlottesville, VA: Evaluation Research Center, University of Virginia, 1977.

TADS

500 NCNB Plaza
Chapel Hill, NC
27514

ROBERT W. COVERT is Associate Professor at the Evaluation Research Center, University of Virginia.
TANYA M. SUAREZ is Associate Director of TADS.

Principal Investigator: Pascal L. Trohanis
SEP Project Officer: Gary Lambour
Managing Editor: Daniel Assael
Typist: Terry Rothra
Editorial Assistant: Iqlia Grennan

This TADScript (Number 7) was prepared for Demonstration and Outreach Projects, State Implementation Grants, and Research Institutes of the Handicapped Children's Early Education Program (HCEEP) administered by Special Education Programs (SEP), U.S. Department of Education.

This material is published by the SEP-funded Technical Assistance Development System (TADS), a support system for eastern Demonstration Preschool Projects and State Implementation Grants. TADS provides technical assistance to designated projects and states when it is requested and needed. TADS is located at 500 NCNB Plaza, Chapel Hill, North Carolina 27514. Our phone number is (919) 962-2001.

This TADScript is distributed pursuant to contract number 300-80-0752 from the U.S. Department of Education. Contractees undertaking such projects under government sponsorship are encouraged to express freely their judgment in professional and technical matters. Points of view and opinions, therefore, do not necessarily represent Department of Education position or policy. The enclosed selections are presented for information purposes only; no endorsement or claim of accuracy is made.

September 1982

TADS is a division of the Frank Porter Graham Child Development Center, The University of North Carolina at Chapel Hill

CONSTRUCTING QUESTIONNAIRES FOR EARLY CHILDHOOD SPECIAL EDUCATION PROGRAMS

The purpose of this TADScript is to provide the reader with a set of guidelines and criteria for constructing a questionnaire. Questionnaires are needed for collecting data and eliciting responses from people about their feelings, needs, attitudes, and behaviors toward themselves or other objects. As unique information-gathering devices, questionnaires hold certain advantages:

- A wide, large group of persons can be surveyed for low cost and time.
- Most questionnaires are completed anonymously, so answers are honest and objective.
- The respondents may take as much time as they wish to fill out the questionnaire and thereby give well thought-out answers.
- Results are uniform because all persons respond to the same questions.
- This uniformity usually provides results which are easy to summarize and interpret.

Let's consider the disadvantages:

- Most often, some questionnaires are not returned. Consequently, the quantity and quality of information diminishes. As the percentage of non-returns increases, the accuracy of the results decreases.
- The validity of the response on a questionnaire is often uncertain. We know nothing of the attitude of the respondents when they fill out a questionnaire and therefore do not know if their responses are colored by extreme interest, complacency or hostility. We also cannot be certain that the person who received the questionnaire was the person who completed the form.
- Another major disadvantage is misinterpretation of the question, leading to an inaccurate response. This problem can be a result of inadequately worded questions or respondents who are not capable of answering the questions.
- A questionnaire's uniformity does not permit the flexibility to explore and gather additional information as do methods such as interviews. The amount and quality of the information received must be based solely on the questions asked.

In spite of these disadvantages, the questionnaire remains a powerful and useful tool for gathering information. You may use it to identify staff development needs, to obtain feedback after workshops and demonstration visits, or to learn about parental opinion of the quality of children's programs. Please remember, however, that the power and usefulness of this tool are related directly to the quality of care you give to the questionnaire's contruction.

In the remainder of the TADScript we will provide you with guidelines and criteria for constructing questionnaires, followed by explanations and examples of good and bad practices. The seven sections of this paper represent areas you should consider carefully as you design your questionnaire. The sections:

I. Focus

II. Title

III. Purpose Statement

IV. Directions

V. Demographic Section

VI. Set of Items and Responses

VII. Structure and Format

To enhance learning, we advise you to review Figure 1, a questionnaire designed to gather information from the parents of children participating in a Handicapped Children's Early Education Program (HCEEP) project (Musumeci & Koen, 1980). Portions of the instrument are referred to throughout the paper. Later, as you read the discussions presented in the paper, your understanding of the writing process should enable you to complete your own questionnaire. To make sure your completed questionnaire is adequate, compare it to information presented in Figure 6 (page 56), a list of guidelines and criteria summarized from the text of this book.

Figure 1. PRESCHOOL HANDICAPPED PROGRAM
PARENT QUESTIONNAIRE

Center For Resource Management, Inc. is conducting an evaluation of the BOCES Preschool Handicapped Program. The basic purpose of the evaluation is to provide a description of how the program is functioning--its accomplishments, constraints and concerns.

Your cooperation and participation with this evaluation effort will give us the opportunity to obtain valuable information from persons involved with program services.

The attached instrument was designed to document your reactions to and perceptions of the Preschool Program. All responses will be held in strict confidence, and only summarized data will be presented in subsequent reports. Your signature is optional, but please supply the following information which will be used for comparison purposes.

1. When was your child first enrolled in the Preschool Program? _____ (Month) _____ (Year)

2. In which program was your child first enrolled? (Check One)
   _____ Classroom Program
   _____ Home Program

3. In which program is your child presently enrolled? (Check One)
   _____ Classroom Program
   _____ Home Program

4. If your child is in the classroom program, please supply the following information:
   a) Name of teacher: _____
   b) Session child attends: _____ (Morning) _____ (Afternoon)

5. Were you ever a parent volunteer in the classroom program? (Check One) _____ Yes _____ No

SECTION I: Involvement in the Program

The Preschool Handicapped Program (PSHP) has provided various means for parents to become aware of and involved in the program over the past year. This section of the questionnaire addresses itself to the parent involvement activities of the PSHP.

1. From what source did you first learn about the Preschool Handicapped Program? (Check One)

Friend/Other PSHP Parent _____
Nursery School _____
Newspaper/Radio _____
Social Service Agency _____
Physician _____
Posters/Flyers _____
Other _____ (specify)

2. In the past year, did you participate in any of the following parent meetings/activities? (Check all that apply)

Child Assessment Conferences _____
Classroom Visits/Observations _____
Screenings _____
Parent Group Meetings (Number attended: _____) _____
IEP Conferences _____
Other Conferences _____
Classroom Volunteering _____
Volunteer Training Program _____
Other _____ (specify)

3. How effective were these parent meetings/activities in:

(Place a check mark in the appropriate box for each item listed below)

	LEVEL OF EFFECTIVENESS				
	Very	Fairly	Somewhat	Hardly	Not At All
a. helping you understand the Preschool Handicapped Program					
b. helping you understand your child's handicap					
c. improving your attitudes about your child's handicap					
d. increasing your skills in working with your child					
e. giving you moral support as a parent					

SECTION II: Reactions to Program

A. Knowledge

Directions: *Please indicate your level of understanding about specific aspects of the Preschool Handicapped Program by circling one number from 1 (low) to 5 (high) for each of the following items.*

		LEVEL OF UNDERSTANDING				
		Low				High
1.	Philosophy of the program	1	2	3	4	5
2.	Goals/purposes of the program	1	2	3	4	5
3.	Screening/placement procedures	1	2	3	4	5
4.	Transdisciplinary team assessment	1	2	3	4	5
5.	Teaching methods of program	1	2	3	4	5
6.	Methods for teaching child at home	1	2	3	4	5
7.	Methods for managing child's behavior	1	2	3	4	5
8.	Legal rights as parents	1	2	3	4	5
9.	Educational rights of child	1	2	3	4	5
10.	Handicapping condition of child	1	2	3	4	5
11.	Services available in community	1	2	3	4	5

B. Attitudes

Directions: *Please indicate your level of satisfaction with the Preschool Handicapped Program by circling one number from 1 (low) to 5 (high) for each of the following items.*

		LEVEL OF SATISFACTION				
		Low				High
1.	Preschool Handicapped Program in general	1	2	3	4	5
2.	Instructional methods used	1	2	3	4	5
3.	Effectiveness of staff	1	2	3	4	5
4.	Frequency of contact with teachers	1	2	3	4	5
5.	Materials used	1	2	3	4	5
6.	Your involvement with the program	1	2	3	4	5
7.	Opportunities for your suggestions	1	2	3	4	5
8.	Screening/placement procedures	1	2	3	4	5
9.	Methods of monitoring child's progress	1	2	3	4	5
10.	Accomplishments of program	1	2	3	4	5

C. Perceived Change in Child

Directions: *The Preschool Program is designed to develop your child's skills in the areas listed below. For each area, please check the level of progress your child has made by circling one number from 1 (low) to 5 (high).*

		LEVEL OF PROGRESS				
		Low				High
1.	Speech and language skills	1	2	3	4	5
2.	Physical and motor skills	1	2	3	4	5
3.	Social skills	1	2	3	4	5
4.	Self-help skills	1	2	3	4	5

SECTION III: Open-Ended Questions

1. What do you see as the major strengths of the Preschool Handicapped Program?

2. What do you see as the major weaknesses of the Preschool Handicapped Program?

3. Would you recommend any changes in the Program? If so, what changes and why?

4. Would you recommend this program to another parent of a preschool handicapped child? _____ Yes _____ No

THANK YOU FOR YOUR COOPERATION!

_____
Parent Signature

I. FOCUS

Certain decisions, if made before the questionnaire is developed, will make the construction of the instrument easier, improve the responses that are obtained, and enhance the results that ultimately are available.

A. The specific purpose(s) of the questionnaire should be determined. Questionnaires are used to gather different kinds of information. Therefore, before questions can be developed, the designer should determine whether to collect facts; test the knowledge of the respondent; discover respondents' opinions, beliefs, or attitudes; or any one of these in combination. The questionnaire used as an illustration in Figure 1 was designed to gather information from parents for the project's evaluation. This definition of purpose, however, does not provide sufficient detail to help understand or construct the instrument. A more specific purpose of the questionnaire might be "to gather information from parents regarding their involvement in the project and their reactions to the project's services for themselves and their children."

B. Identify the persons who will use the results and identify their reasons for wanting them. Identifying the individuals or groups who will use the results of your questionnaire may help determine the purpose of the questionnaire and guide your choice of items. If, for example, the project staff will be using the results to improve their efforts, you could ask questions about specific components of the project. On the other hand, if a governing or funding agency will be using the results, questions regarding the overall impact of the project might be asked.

C. Identify the major categories of information you seek. Identifying categories of information needed means clarifying the content of the questions that will be developed. A content outline will help you define the purposes and scope of your instrument. For our sample instrument, the content outline might be:

Content Outline
Parent Questionnaire

A. Demographic Information

1. Length of time child has been in the program
2. Type of services in which the child was enrolled
3. Parent volunteer activities

B. Parent Involvement

1. Source of information regarding the program
2. Type of involvement in the program
3. Reactions to parent activities

C. Services for Children

1. Knowledge of program for children
2. Satisfaction with program for children
3. Perceived change in child

D. General Opinions

1. Strengths
2. Weaknesses
3. Recommendations

After the instrument has been developed, the items should be compared with the content outline to see if all areas have been included.

D. Identify the specific group(s) that will answer the questionnaire. Be specific so that: (a) the targeted groups will receive the questionnaire; (b) the results can be described as representing a particular group; and (c) the characteristics of the group can be taken into account when you construct the instrument. The questionnaire in Figure 1 obviously will be sent to parents. Further specifics might also be helpful -- e.g., only

to mothers, only to the parent who cares for the child most of the time, etc.

E. Identify the type of information you seek. Preliminary decisions regarding the most useful type of information will help you later to select appropriate items. Some questions (such as those regarding length of time the child has been in the program) might best be answered with actual numbers. If you wnat to obtain opinions that can be compared, use a standard item with specified responses. Finally, if you desire individual opinions, use questions which permit a variety of responses.

When these preconstruction decisions have been made, a great deal of information will be available for actual construction of the instrument.

II. TITLE

Every instrument requires a title to summarize briefly the content of the instrument. Three criteria should be considered:

A. The title should reflect the content of the instrument. This can be done in at least two ways. First, the title simply can reflect the content of the items. For example, if a person has constructed a questionnaire on attitudes toward the handicapped, an appropriate title might be: "Survey of Attitudes Regarding the Handicapped." A title also might reference the content of the items and the major type of item in the instrument. For example, if you develop an instrument to assess a teacher's ability to teach, and all or most of the items are 5-point rating scale items, you might appropriately title your instrument a "Teacher Rating Scale." It is important that the title does not mislead the respondent.

A title also may summarize the purpose of the instrument: "Parent Evaluation of Child Services."

B. The title should be concise and to the point. Long titles tend to distract and divert the respondent from the task of actually filling out the questionnaire. The title should allow the respondent to grasp the scope of the instrument quickly without becoming bogged down with specifics. The titles above are good examples. A wordy and, consequently, bad title for a political attitude survey would be: "A Survey of Attitudes, Reactions, and Observations of Legislators Regarding the Handicapped in the U.S. During the Past Ten Years."

C. The title should be written in plain English. Complex or unrecognizable words or phrases in the title of the instrument might discourage potential

respondents, so eliminate jargon and acronyms. In general the title should be easy to understand and informative. Your well-designed instrument may remain unanswered if people are alienated or confused by the language in the title. Writing the title is usually best done after construction of your instrument, so that it more accurately reflects both content and/or method of eliciting responses.

III. INTRODUCTORY STATEMENT

An introductory statement should contain information about the instrument's purpose and how the data will be used, anonymity, and motivators for respondents to complete the questionnaire. This introductory statement usually is found in one or two brief paragraphs following the title of the instrument, or in an accompanying cover letter.

Consider the following:

A. The introductory statement should include a brief summary of the instrument's purpose. This statement indicates how and by whom the information will be used and tells respondents the kind of project with which they are involving themselves. The statement should be brief and straightforward, with no intentional or unintentional bias.

Looking back to Figure 1, the introductory paragraph states that the purpose of the instrument is to contribute to the evaluation of a program for preschool handicapped children. More specifically, the information gathered will be used to describe how the program is functioning -- its accomplishments, constraints, and concerns.

As mentioned previously, the purpose and use of the questionnaire should be written before actually constructing the instrument. The purposes, if more than one exist, should be consistent. Inconsistency can cause respondents to lean toward noncommital responses, which result in data that contain little real information. The statement of purpose also reminds the instrument constructor to develop items within a certain scope --i.e., information pertinent to the use of the instrument. A side benefit of providing the respondent with a description of how the

information will be used is that it provides each respondent an opportunity to gauge the face validity[1] of the instrument.

B. The introductory statement should include an appropriate statement concerning the respondents' anonymity. This statement should make explicit whether or not an individual's completed instrument will be used and be made accessible to others or only as part of an aggregation. Obviously, divulging respondents' negative comments or suggestions might have serious repercussions for them and thereby temper their responses. Furthermore, it is illegal to elicit information under pretext and use the information for another purpose. In Figure 1, the introduction states, "All responses will be held in strict confidence, and only summarized data will be presented in subsequent reports." Exclusion of this statement might cause respondents to think that their comments, individually identified, may affect services to them or their child.

C. The introductory statement should use a "motivator" to encourage the respondent. Examples of frequently used motivators include promising respondents a copy of the results of the survey and making respondents feel as though they will have an impact on an organization. More expensive motivators include free tickets, money, and access to generally inaccessible places (i.e., reserved parking areas or discounts on purchases). A well-known gardening magazine recently promised respondents a free plant if a 200-item questionnaire was completed. Often, outside motivators exist which ensure replies --e.g., withholding a person's paycheck or making the response a general expectation.

---

[1]Face validity is the extent to which a respondent judges the instrument to measure what it purports to measure.

Motivators take a variety of forms. The important considerations are: first, can you afford the motivators, and second, does the pressure of the motivator potentially bias the responses? If a motivator works, more respondents reply, and the data gathered usually will better represent the sample or population surveyed.[2]

D. The introductory statement should be written at the language level of the respondents. Here, as in the title, and in all other sections of an instrument, the language used should be understood easily by the respondents. This seems obvious, but remember, if you do not know the reading level of the respondents, you may unwittingly write items at your own level. The problem can easily be remedied by administering the questionnaire to a subsample of respondents and questioning them verbally to check their understanding of the terms and expressions. The instrument designer and the respondent must share these meanings or the instrument will not measure what it purports to measure.

One last word concerning whether introductory statements are required in all instruments. In many situations the factors listed above are implied. If any of these areas are not clearly implied or stated elsewhere, a statement pertaining to each should be included.

---

[2]This is not true if the people responding consistently represent one portion of the population rather than another. For example, if your population potentially had an equal number of Republicans and Democrats, and Democrats consistently filled out the questionnaire more than Republicans, your overall results would contain that bias. Increasing the sample responding would not necessarily eliminate the bias.

IV. DIRECTIONS

Basically, directions come under two headings: general directions relate to the total questionnaire; specific directions deal with individual sections of the instrument. At this point, only general directions will be considered (specific directions will be explained in other sections of this book).

Consider these important factors when you write general directions for your questionnaire:

A. Directions should be complete, unambiguous, and brief.

B. Directions should be written at the language level of the respondents.

C. Directions should tell respondents to whom they should give the instrument once they have completed it. In some situations the persons who administer the instrument may hand out the questionnaire and tell the respondents to return it when it has been completed. In this case written directions are not required. If, on the other hand, the respondent receives the questionnaire in the mail, clear instructions are necessary. A self-addressed, stamped envelope facilitates easy return. This procedure adds to the cost of the survey, but it generally increases the response rate. Further, it is a good idea to give the respondent a deadline to complete the instrument -- e.g., "Please return the completed questionnaire within the next week." This time specification can make the respondent aware of the importance of returning the questionnaire immediately. Another example of a time specification might be: "Only those responses received by October 1 will be used in the data summaries." A clear time deadline particularly is effective when respondents are motivated to

answer a questionnaire because the results of the instrument directly affect them.

D. When the questionnaire is accompanied by a separate answer sheet, the general directions should specify how the respondent should fill out the answer sheet.

E. The general directions should explain to the respondent how to deal with items which are not applicable. Does the respondent leave those answers blank? Is an appropriate space provided? Is the respondent expected to respond to all items? Ensure that all respondents respond in a similar way to the nonapplicable items, and the results will be easy to interpret and analyze.

F. The general directions should specify the approximate amount of time required to complete the instrument. This criterion is particularly applicable for a long questionnaire. Time approximations can help alleviate a respondent's concern that the questionnaire will take too long to complete.

As is the case with the introduction, place the general directions according to the structure of the questionnaire. The beginning of the instrument or in the cover letter are two appropriate places.

All the criteria stated above are not required for every instrument, but they all should be considered to determine which are appropriate for a given questionnaire.

V. DEMOGRAPHIC SECTION

In almost every situation, whether a person is developing a questionnaire for an evaluation study or to gather input, it is useful to obtain certain demographic information. For example, in Figure 1, the demographic variables are found on the first and second pages. They ask for information regarding the services provided to children by the program and the parent services in which the respondent participated.

Though demographic variables differ according to the type and purpose of the questionnaire, some general guidelines can be followed.

A. The demographic information should be limited to only those variables that will be used to answer specific questions. All of us are familiar with questionnaires that begin with lengthy demographic sections which elicit information that seems unrelated to the overall purpose of the questionnaire. To avoid this problem, the person constructing the instrument should consider how the information gained from any given demographic item will be used in data analysis. If uses for an item -- in particular or in comparison with other items -- cannot be identified, then it should not be included.

Demographic information which will help you understand the data is important. To select appropriate demographic information, review a list of standard characteristics for the insights they may give to the data. The list might include: age, race, sex, and type of handicapping condition of the child -- just to mention a few. Identify the critical information you need and make sure you get data on only those variables.

B. The demographic section should reveal the relationship of the respondent to the object of measurement. One important category of demographic data sometimes is overlooked and relates to the respondent's judgments about another person's performance. For example, parents are often asked questions about the project staff, or project staff are asked questions about other service providers or agencies. In these situations, the respondent must divulge the amount of contact with the person or thing being rated. If a parent is rating a center-based teacher, for example, determine the extent to which the parent has observed the teacher in the center. Certainly, data from a parent who has observed a teacher should be viewed differently from data given by a parent who has not observed a teacher.

C. Certain items in the demographic sections should be optional. Though this criterion rarely is important, it has specific implications for situations where decisions about individual respondents will be made. For example, if you are polling potential applicants for an advertised position, federal law prohibits considerations based on sex, race, or religious preference. This problem can be avoided by briefly explaining why the information is needed and requesting that the respondent provide answers to these questions.

D. Language in the demographic section should be appropriate to the language level of the respondents.

E. Demographic items should be grouped together. The positioning of demographic items is not important, but it is usually a good idea to group them together. In Figure 1, the demographic items come at the beginning of the instrument. In most cases, the demographic section follows the introduc-

tory statement and the general directions, though sometimes it is located at the end of the questionnaire.

Sometimes demographic information may be left out of the questionnaire because that information can be obtained elsewhere. For example, if you were conducting a survey of the incidence of handicapped young children in the community, you probably would not have to ask each respondent the incidence of low-birth-weight babies they encounter; this information could be obtained easily from some other source such as hospital records or birth certificates. Obtaining appropriate demographic information from other sources accomplishes at least two purposes: first, you relieve the respondents of a somewhat tedious task; second, you ensure the compatability of the demographic data across your respondents. Another illustration of this procedure might involve a survey of children who have left a program. If the program has kept an up-to-date file on the children, demographic information could be found in the file rather than from the respondents.

VI. WRITING ITEMS

Types of questionnaire items include: rating scales, checklists, short answers, yes-no questions, open-ended questions, ranking questions, and several others. Each method presupposes certain assumptions, uses, advantages, disadvantages, construction, directions, and data analysis techniques. This section will discuss all such points of interest. If you have not selected your item type or wish to confirm your selection, read the entire section for help. If you have chosen your item type already, read the discussion which applies to that type of item.

Before describing specific types of items, several comments are in order. Two parts are inherent in any item type: the stem elicits the information; the response is what the respondent does to complete an item. Generally, the responses are supplied freely by the respondent or are selected by the respondent from a list provided by the instrument designer. In the construction of any instrument, various types of items can and should be used, depending on the information to be gathered and the eventual use of data. The developer should, therefore, consider the best way to obtain the required information and use as many of the various item types as are appropriate.

A. Rating scales. A rating scale is a device to indicate the degree to which a certain trait or attitude is present in a person, place, or thing. A rating scale is made up of two parts: a stem and a response set. The stem is the stimulus which elicits the response. The respondent selects an answer from the response set which contains categories which may or may not be numerical.

The following assumptions are made when one uses a rating scale:

** The attribute can be scaled along some continuum. For example, an item rating the effectiveness of a training workshop assumes that the attribute "effectiveness" can be scaled from ineffective to highly effective.

** The respondent is presented with a continuum that is isomorphic to his/her internal continuum. In other words, the continuum applicable to the actual attribute is consistent with that within his/her self. Perhaps the best example of this would be one in which the trait exists but there is no recognition of it. If the particular trait to be rated were greenness and the ratings were to be made on a scale between blue and yellow, then a colorblind person would not have an isomorphic continuum on which to rate the particular attribute.

** The respondent can consistently discriminate between different levels of the continuum.

Empirical testing of these assumptions would be taxing to say the least. Importantly, however, each potential constructor should consider whether or not these assumptions are pertinent. If you cannot make the assumptions listed above, employ alternative item types.

This paper will discuss three of several formats for writing ratings scales: numerical (currently the most popular), graphic, and Likert.

A numerical scale requires the user to indicate a particular number in a stem which may or may not be associated with verbal description. Figure 2 shows several different formats for numerical scales.

These examples represent only a small variety of numerical rating scales. The items in Figure 1 are additional illustrations. Though numerals do not appear in most cases, numerical scales will produce ordinal or nominal data. The choice of one of these types depends largely on the preferences of the instrument designer --who has knowledge of the respondents and the total structure of the instrument.

Graphic scales present a stem along with a response set containing a line on which the user designates a response. Examples of the graphic type of scale might look like this:

Figure 2

Examples of Numerical Rating Scales

Example: Assume that you have been instructed to develop an instrument to determine parents' perceptions of a parent/child activity guide you have developed. One of the variables of interest may be the parents' perceptions of the readability of the guide.

1. Rate the readability of the activity guide compared with other guides for parent/child activities you have read.

( )	( )	( )	( )	( )
very difficult to read	difficult to read	about average to read	easy to read	very easy to read

2. Rate the readability of the activity guide compared with other guides for parent/child activities you have read.

( )	( )	( )	( )	( )
1	2	3	4	5

3. Rate the readability of the activity guide compared with other guides for parent/child activities you have read.

( )	( )	( )	( )	( )
very difficult to read		about average to read		very easy to read

4. Rate the readability of the activity guide compared with other guides for parent/child activities you have read.

_____a. very difficult to read
_____b. difficult to read
_____c. about average to read
_____d. easy to read
_____e. very easy to read

5. Rate the readability of the activity guide compared with other guides for parent/child activities you have read.

☐	☐	☐	☐	☐
very difficult to read				very easy to read

6. Rate the readability of the activity guide compared with other guides for parent/child activities you have read.

_____	_____	_____	_____	_____
very difficult to read				very easy to read

Rate the readability of the activity guide compared with other guides for parent/child activities you have read.

very difficult to read	difficult to read	about average	easy to read	very easy to read

Rate the readability of the activity guide compared with other guides for parent/child activities you have read.

very difficult		about average		very easy

Likert-type scales are somewhat different from numerical or graphic scales -- each set of responses, ranging from "strongly disagree" to "strongly agree," is invariant and not specific to the stem. As an example, the Likert-type item would require a modification of the stem in order to be comparable to the previous example:

The activity guide is very easy to read compared with other guides for parent/child activities I have read.

( ) strongly disagree	( ) disagree	( ) undecided	( ) agree	( ) strongly agree

The activity guide is very easy to read compared with other guides for parent/child activities I have read.

( ) SD	( ) D	( ) U	( ) A	( ) SA

Obviously, these examples show a variety of rating scales which elicit the same information. When you have made the necessary assumptions and have chosen rating scales, consider these criteria:

1. <u>The stems of rating scale items should be written unidimensionally.</u> This means that only one attribute or trait should be described in the stem. In the first two sets of examples, the attribute to be rated was readability. Furthermore, readability was limited to comparison of this guide with others. This eliminates the possibility that a respondent will compare the readability of the activity guide to other types of reading matter. It also allows the person reviewing the data to make a straightforward interpretation. Here is an example of an item which is not unidimensional:

> Rate the effectiveness of the activity guide and the assistance provided by the home teacher in enhancing your child's development.
>
> | ( ) | ( ) | ( ) | ( ) | ( ) |
> |---|---|---|---|---|
> | extremely ineffective | somewhat ineffective | neither effective nor ineffective | somewhat effective | extremely effective |

Interpretation of data from this question would be difficult, since an answer of "neither effective nor ineffective" could have several interpretations:

a. The activity guide was effective while the assistance from the home teacher was ineffective.

b. The assistance from the home teacher was effective while the activity guide was ineffective.

c. Both were neither effective nor ineffective.

Though this and many of the following criteria may seem fairly obvious, the good instrument designer will apply each criterion to every item to ensure that all items are unidimensional. The problem presented in the example above could easily be resolved by splitting

the item -- one question concerning the effectiveness of the activity guide and the other concerning the effectiveness of the assistance provided by the home teacher.

2. The response sets to rating scales should be written unidimensionally. This criterion is related closely to the previous one and relates to the the response set of any item. To ensure unidimensional meaning in the response set, refer only to the attribute described in the stem. In our previous example, we see that though the stem had two things to be rated, the response set was unidimensional (i.e., effectiveness). Criteria 1 and 2 should be considered jointly for each rating scale item employed.

3. The response set of a rating scale should be tied logically to the stem. For example, if you ask the respondent to rate the readability of a guide, the following response set would be inconsistent with the stem.

( )	( )	( )	( )	( )
definitely	probably	don't know	probably not	definitely not

4. The level of specificity of the stem item should answer the user's needs. For example, if the purpose of an item is to determine generally whether one teacher is more effective than another, the question could be phrased, "Rate the overall effectiveness of the instructor." If the purpose is to provide feedback to individual teachers concerning their various teaching abilities, a general item would not be particularly useful. A series of rating scales assessing specific

teaching acts -- ability to plan and work individually with children or lead play -- might prove to be more informative.

5. Use the rating scale item only when a more direct method is not available. Rating scales generally are overused, and the extent to which their assumptions can be met is difficult to document. Therefore, rating scales should be used only as a last resort or with more reliable and valid item types. Section 3 (previous page), Readability of an Activity Guide, represents a good example of the misuse of rating scales. In our examples of items rating readability, analyzing the words used in the activity guide by grade level and comparing that with the actual reading level of the participants is a more direct way of getting the same information.

6. Directions for using the rating scale should be included, along with appropriate examples. Unfortunately, many questionnaires ask for ratings along a certain continuum, but do not specify how those ratings should be made. For example:

I hate it /_____/ /_____/ /_____/ /_____/ /_____/ I love it

Here, the actual rating task is unclear. Should you place an X between the lines? If so, what does an X in the middle space mean? Should you circle one of the oblique lines? Obviously, the designer of the instrument had a clear idea of the appropriate response but did not make that clear to the respondent. To ensure that all respondents know how to respond appropriately, the designer should (1) provide explicit directions of rating procedures and (2) provide an example. Let us look again at the examples in the introduction to this section.

Obviously, slightly different directions are in order for each individual case.

a. Directions for a Numerical Rating Scale

You are to rate the following items by placing a checkmark in the appropriate space under each item. For example, if you feel that the guide is as readable as activity guides for parents and children, you would place a check in the space labeled "about average to read," as shown below.

Rate the readability of the activity guide compared with other guides for parent/child activities you have read.

( )	( )	(✓)	( )	( )
very difficult to read	difficult to read	about average to read	easy to read	very easy to read

b. Directions for a Graphic Rating Scale

1. In the following set of items, you are to rate each item by placing an "x" over the set of words which best describes your feelings. For example, if you feel the activity guide is more readable than guides for parent/child activities, your response would look like this:

   Rate the readability of the activity guide compared with other guides for parent/child activities you have read.

				X
very difficult to read	difficult to read	about average to read	easy to read	very easy to read

2. In the following set of items, you are to rate each item by placing a check in the appropriate space. As you will note, only the ends of the scale are marked. Your task, therefore, is to consider each item on a five-point continuum, decide which line on the continuum corresponds with your feelings, and place an "X" on that line. For example, if you feel that the activity guide is a little more difficult than average your response would look like this:

   Rate the readability of the activity guide compared with other guides for parent/child activities you have read.

very difficult to read		✓		very easy to read

c. Directions for a Likert Scale.

In the following set of items, you are to decide whether or not you strongly agree (SA), agree (A), are undecided (U), disagree (D), or strongly disagree (SD) with each of the statements listed, and indicate your preference by circling the appropriate letters. For example, if you agree with the statement that the activity guide is very easy to read, you would circle "A."

The activity guide is very easy to read compared with other guides for parent/child activities I have read.

SD D U A SA

The directions above, though not comprehensive, suggest the types of statements that ensure understanding by the respondent. The importance of these directions should not be underemphasized, since the validity of the data depends on the extent to which the subject understands how to respond. Directions, in general, should be as brief as possible. If you are sure that your respondents are familiar with your rating scale, no examples are necessary, however specific directions for any set of similar items are advisable.

7. The language used in rating scale stems and responses should be written at the respondents' level of understanding.

8. The type of rating scale format selected should be easily understood by all respondents. If a question arises as to whether or not a particular scale is appropriate to a given group, the best procedure is to administer several types of scales to a sample target group. You should question them as to the procedure with which they feel most comfortable and which best allows them to express their opinions. This procedure, though time-consuming, will help ensure valid results.

9. Rating scale items should not elicit biased responses. Item stems and responses should be written in such a way that the user will not think that certain responses are preferred. Example:

> Rate the extent to which you enjoyed the immoral movie "Peyton Place."
>
> | ( ) | ( ) | ( ) | ( ) | ( ) |
> |---|---|---|---|---|
> | one of the worst I have seen | | | | one of the best I have seen |

Clearly, the bias in the stem suggests that a person should rate this item low rather than high. A quick perusal of item stems by a colleague or other qualified person should help eliminate this problem.

10. The response sets for a rating scale should present choices along similar degree continuums. This criterion probably is the most controversial. Many questionnaire writers argue that all positive or negative poles similarly situated on a page encourages "patterned" responses. That is, the respondent simply may mark all items in the same way without thinking about them individually. Our experience indicates that the practice of reversing response poles is very confusing to users and therefore tends to invalidate their responses. Potential bias may be counteracted by structuring the rating scale items in such a way that response sets are specific to only one question and, together with the stem, comprise a single entity (see Figure 3). Each stem and its response set graphically displayed as an entity also is useful when response sets are the same across items. If you can establish that users do not respond similarly to items (by noting previous questionnaires or answers to a short pretest), then choose the

Figure 3. Example of Individual Layout for Rating Scales

Interest in children

( ) always enthusiastic	( )	( ) sometimes enthusiastic	( )	( ) never enthusiastic

Availability to parents

( ) always available	( )	( ) sometimes available	( )	( ) never available

Personal appearance

( ) always well groomed, clothes neat and clean	( )	( ) somewhat untidy at times	( )	( ) slovenly, clothes untidy and unkempt

Stimulates intellectual curiosity

( ) inspires children to independent effort	( )	( ) occasionally inspiring	( )	( ) destroys interest in activity, makes participation repulsive

matrix format (see Figures 4 and 5). The primary advantage of the matrix format is its use of space; maximum information can be obtained in minimal space.

Figure 4. Matrix Layout for Rating Scales

Rate your child's teacher on the following attributes from below average to above average, by circling the appropriate number.

	way below average				way above average
Interest in children	1	2	3	4	5
Availability to parents	1	2	3	4	5
Personal appearance	1	2	3	4	5
Stimulating intellectual curiosity	1	2	3	4	5

Figure 5. A Variety of the Matrix Layout

Rate your child's teacher on each of the attributes below by placing the appropriate number on the space provided.

1 = way below average
2 = below average
3 = average
4 = above average
5 = way above average

_____1. Interest in children

_____2. Availability to parents

_____3. Personal appearance

_____4. Stimulating intellectual curiosity

Another way to avoid response bias is to intersperse negative and positive stems, while keeping the response sets the same. The is appropriate particularly with the Likert-type scales. For example:

This parent/child activity guide is one of the most readable I have encountered.

( )	( )	( )	( )	( )
strongly disagree	disagree	no opinion	agree	strongly agree

This parent/child activity guide is one of the least readable I have encountered.

( )	( )	( )	( )	( )
strongly disagree	disagree	no opinion	agree	strongly agree

11. Avoid using global terms in response sets. Examples of global terms include: best to worst; all of the time to none of the time; and high score to low score. The major problem with global terms is that their interpretation lies with each respondent. For example, if instrument designers were interested in how often the respondents went to the movies, they could ask the question in the following ways:

Rate the extent to which you attend movies.

( )	( )	( )	( )	( )
never attend	infrequently	sometimes	often	regularly attend

A better way to elicit information on this same variable would be:

( )	( )	( )	( )	( )
never	once or twice a year	once a month	once a week	more than once a week

The latter construction helps ensure that every person is using the same continuum of response and providing information which is easily interpreted.

12. Rating scales should include from three to seven categories. To decide on the number of response categories to use for a particular item, consider the trait being rated and the extent to which individual categories can be identified. Ask yourself these questions: Is it possible for the respondent to make the discrimination demanded of individual categories? Should an even or odd number of categories be provided? The advantage of an even number of categories is that the respondent must be either positive or negative about the attribute (not noncommittal). The advantage of an odd number of categories is that the respondent can be noncommittal.

13. Choose the format for your rating scales that corresponds to the sophistication of respondents. Generally, two basic formats exist. A matrix format is shown in Figures 4 and 5, and an individual format is shown in Figure 3. All examples are based on the same item stems.

The matrix arrangement allows the instrument designer to cover more variables in less space. However, it necessitates the use of general terms in the response set, since it must be appropriate for many different stems. Further, it may allow the subject to respond to all items in the same fashion rather than responding to each item individually. For example, if parents liked a teacher, they might rate the teacher high on all items without reading each item specifically. Generally, the matrix display structure should be used only with respondents who are sophisticated in the use of rating scales.

The individual format is effective with a less sophisticated group, since it forces respondents to read each item individually and offers response sets specific to individual items.

14. Allow for a "not applicable" response when appropriate. Respondents may not have had the opportunity to form perceptions about specific traits the questionnaire assesses. So, a "not applicable" response should be available. However, this option can cause problems in later aggregation of data. Use the "not applicable category" only when appropriate: example -- If the presenter in workshop A and B used audiovisual equipment, but none was used in workshop C, then attendees of workshop C should not have to rate the effectiveness of audiovisual materials.

15. Choose an analysis strategy appropriate to the data produced and the uses for which the questionnaire was designed. Once you have completed the construction of your questionnaire and administered it to a set of respondents, you are ready to analyze and report the results. It is always a good idea to summarize or tally the responses before deciding on specific analyses. A careful review of the actual results often will indicate the areas for and types of analyses that would be most useful. Then the data can be analyzed in a variety of ways depending on the assumptions made about the scales themselves. If the designer assumes the scales are of an interval nature, all types of parametric statistics are appropriate (when their assumptions are met). If these assumptions cannot be justified, then nonparametric statistics should be used. Further discussion of these topics can be found in texts on basic statistics.

B. Qualitative selection items. In the preceding section, Rating Scales, we discussed selection of responses based on some continuum relating to the stem. In this section, we will discuss responses that are not on a continuum. For example, if you are interested in the sex of your respondent, you have the stem "sex" and the responses "1, female" and "2, male." Another example is the following:

What is your position in the program?

|_____| Administrator

|_____| Teacher

|_____| Occupational Therapist

|_____| Social Worker

|_____| Other

The purpose of this type of selection item is to elicit qualitative information from the respondents. For example, job classification, sex, or race are more suited to this type of item than number of children, hours of participation in the project, etc. Responses to the former will be qualitatively different, while responses to the latter will be quantitatively different. Qualitative response categories also are used to elicit demographic information; all respondents can select from the same set of response categories. This type of response also is appropriate when all or most of the response categories are known. Finally, this type of item requires no judgment in the coding of responses, since all interpretative judgments are made when the instrument is constructed.

At least six rules should be applied when constructing items of this type.

1. State clearly and unambiguously the stem and responses.

2. Use language appropriate to the respondents.

3. State the stem unidimensionally.

4. Write an exhaustive response set. When you write responses to an item, include all possible choices. For example, if you are surveying the positions of the respondents, you may include the following selection items:

Identify your position.

|_____| Administrator
|_____| Teacher
|_____| Occupational Therapist

You have not been exhaustive, since there is no place for the social worker or teacher to check. This problem can and should be avoided by including an "other" or "miscellaneous" category, as demonstrated below.

Identify your position.

|_____| Administrator
|_____| Teacher
|_____| Occupational Therapist
|_____| Other (please specify)

The "other" category makes the item applicable to all respondents and finds out who or what the other categories include. This latter benefit is of particular significance both in qualifying your results or in future construction of similar items.

5. The response categories should be mutually exclusive. Mutual exclusiveness means that the categories do not overlap. Numerical categories most frequently overlap. An example:

> How much time do you spend working with your child on project activities during the average week? (check one)
>
> |_____| 0 - 1 hours
>
> |_____| 1 - 3 hours
>
> |_____| 3 or more hours

The parent who estimates either 1 hour or 3 hours must choose a category arbitrarily. Several alternatives are available to ensure mutual exclusiveness in responses for numerical data. First, don't use selection-type items when numerical categories are the required responses. You may use instead an open-ended type of item (to be discussed more fully in a later section). Second, if numerical categories are to be used, ensure mutual exclusiveness by naming limits that do not overlap. The above item would then read:

> How much time do you spend working with your child on project activities during the week? (check one)
>
> |_____| Less than one hour
>
> |_____| Between 1 and 3 hours
>
> |_____| More than 3 hours

6. Supply direction for difficult items. Two basic types of directions can be used: those which force respondents to choose only one of the

response categories; those which allow respondents to select all responses that are appropriate. The type of directions you use depends on the purpose of your question. Suppose that as a program administrator you are concerned with the sources of information the teachers use to develop IEPs. Your question might read:

> Identify the information you use to develop IEPs (check all that apply).
>
> ____ standardized assessment results
>
> ____ parent reports
>
> ____ observation
>
> ____ curriculum assessment tools
>
> ____ other, please specify _____

This question indicates all the information that is used and the extent of use of each response. If you wanted to know which source was used most extensively, the directions might state: "Check the <u>one</u> source you most often use." Directions also ensure consistent answering procedures. Without specific directions, some respondents may check all of the appropriate choices, and others may only check the most appropriate. The data then might be weighted or biased.

Another selection item requires the respondent to rank responses or qualitative statements. To rank a response, the respondent orders the qualitative statements according to criteria stated in the stem. If you are trying to ascertain parents' preferences for parent meetings or activities, for example, a ranking item might look like this:

Rank the following activities from 1 to 5: 5 being the one in which you most like to participate, 4 being the one you would next prefer, and so forth. Each activity should be ranked only once, and no number should be used more than one time.

____ classroom visits/observations

____ parent group meetings

____ IEP conferences

____ Volunteer Training Program

____ Other (please specify) _____

Data from ranked items can be analyzed to produce results very similar to results from qualitative items, since frequency of response and percentage are the basic types of analyses associated with all qualitative items. The data also can be used in the same way as demographic data. For example, by dividing a group of respondents according to their response on item "age of child," you can determine if parents of younger children answer certain questions differently than parents of older children.

C. Supply items. This type of item differs radically from items previously discussed. In a supply item, the respondent must produce the response to an item rather than select from a set provided in the questionnaire. Supply items are stems only. This type of item is particularly effective when the domain of responses is unknown, or when in-depth information about a subject is sought, such as a person's feelings, suggestions, or an explanation of events. The data gathered from these types of items can render a rich description of happenings and can provide very detailed data because the respondent is not limited to general response categories of selection items. Basic disadvantages are two: first, it is hard to get

responses to open-ended questions; second, aggregating the data is extremely time-consuming.

1. Whenever possible, appropriate units should be specified for supply items. As noted previously, respondent-specific information is one of the most appropriate uses of the supply item. For example, you might be interested in the average number of hours spent studying per week, height and weight, hours spent working, monthly or weekly wages, etc. When you write items that elicit this type of information, tell respondents the unit of measurment with which they should answer. For example, if you are surveying parents to determine how much time they work with their children on project-related activities, you could ask the following open-ended question:

> How much time do you spend working with your child on project-related activities?

The question, as it stands, would allow all sorts of responses, including: "about 10 hours per week," "only on the weekends and full time during the summer," "about 2 hours per day," or "one week out of four." If you want answers comparable across respondents, such as relating the amount of time working to the developmental progress of children, it is essential that the item stem require more specific answers. A better item might be:

> During the time that your child was enrolled in the program, what was the average number of hours per week that you spent working at home with your child on project-related activities?
>
> _____ hours per week

Another example reveals a different aspect of the problem. In a recent survey of college administrators concerning volunteer activities on their campus, a question was asked about the amount of money allocated at their institutions for volunteer activities. A perusal of these results revealed the following responses:

less than $0

$866

$4,000

$25

around $10,000

How is it possible to reasonably aggregate these results, since the degree of accuracy varies so greatly? We assume that the person who said "$866" was accurate to the nearest dollar, though this probably is not true of the person who replied "around $10,000." Some administrators answered to the nearest dollar, while others answered to the nearest thousand dollars. If the purpose of this question is to aggregate data across institutions, it is necessary to include not only the units (in this case dollars), but also several examples. The question should read:

> Please specify to the nearest one hundred dollars the amount of money which is budgeted to your volunteer program. For example: if your budget is between 0 and $49, enter 0; if your budget is between $50 and $149, enter $100; if your budget exceeds $1,000, please report it to the nearest one hundred dollars.

2. Provide sufficient space for responses. Nothing is more frustrating than to receive a questionnaire that does not leave enough space for your responses. In general, designers should consider the type of

responses they are looking for and should provide space accordingly. Remember, handwriting varies; if lines are provided, don't put them too close together. If you expect respondents to write longer responses to some questions than to others, allocate space accordingly.

Before leaving the discussion of the various item types and their criteria, one general statement should be made concerning their individual uses in a questionnaire. When designing any type of questionnaire for whatever purpose, select those item types which elicit the most reliable and valid responses concerning the variables addressed. That means that for any instrument, do not limit the format to one type of item; rather, vary the format as appropriate.

VII. STRUCTURE AND FORMAT

After writing your items, directions, and demographic section, put the questionnaire together in a short, clear, final form. Several criteria can be applied at this phase of the instrument's construction.

A. Items should be grouped according to item types or similarity of content. A logically structured questionnaire is easy to fill out and therefore helps increase the response rate. Your choice of format should depend on the purpose, the variables, and the item type selected. When long lists of items are used, it is sometimes appropriate to set the groups of items apart as subsections of the instrument. Once again, the purpose is to simplify the task of filling out the questionnaire. If you categorize item types into subsections, write directions for each subsection.

B. Group items within sections according to ease with which they can be answered. If your instrument is made up of both selection and supply type items, the selection type should precede the supply types, since a greater number of respondents are likely to check rating scales than they are to fill in open-ended questions. This criteria helps ensure that respondents do not get discouraged by the more difficult items before they have answered the simpler ones.

C. Length of the instrument should relate to respondents and purpose. Instrument designers usually face a dilemma when they decide how long to make the instrument. They want to get as much information as possible, while at the same time obtaining the greatest response rate. Generally, short questionnaires get higher response rates than long questionnaires. When you decide on the length, determine your respondents' tolerance level

for answering questions and balance that level against the minimum amount of information you require. The kind of motivator provided for the respondent also should affect the length of the instrument. Highly motivated respondents will answer more questions than respondents who have no incentive to fill out the instrument. Finally, never ask questions whose answers you have neither the time nor ability to analyze.

D. The instrument should be clearly reproduced. The impact of a well-written questionnaire can be negated simply by typographic errors, or blurred or unclear copy. A respondent will not be inclined to take seriously a messy questionnaire. Photocopies generally are clearer than mimeographs, and the former make the instrument look more professional. Finally, nothing takes the place of a good proofreading.

SUMMARY

When you develop or critique any instrument, refer to Figure 6 (next page) to see all the criteria explained throughout this book. This list can be used as a guideline in the development process or as a critical tool by separately applying each of the criteria to your instrument. note that outline numbers and letters in Figure 6 are keyed to appropriate discussions within the text of this book.

Figure 6. Guidelines and Criteria for Constructing Questionnaires

I. FOCUS

A. The specific purpose(s) of the questionnaire should be determined.
B. Identify the persons who will use the results and identify their reasons for wanting them.
C. Identify the major categories of information you seek.
D. Identify the specific group(s) that will answer the questionnaire.
E. Identify the type of information you seek.

II. TITLE

A. The title should reflect the content of the instrument.
B. The title should be concise and to the point.
C. The title should be written in plain English.

III. INTRODUCTORY STATEMENT

A. The introductory statement should include a brief summary of the instrument's purpose.
B. The introductory statement should include an appropriate statement concerning the respondents' anonymity.
C. The introductory statement should use a "motivator" to encourage the respondent.
D. The introductory statement should be written at the language level of the respondents.

IV. DIRECTIONS

A. Directions should be complete, unambiguous, and brief.
B. Directions should be written at the language level of the respondents.
C. Directions should tell respondents to whom they should give the instrument once they have completed it.
D. When the questionnaire is accompanies by a separate answer sheet, the general directions should specify how the respondent should fill out the answer sheet.
E. The general directions should explain to the respondent how to deal with items which are not applicable.
F. The general directions should specify the approximate amount of time required to complete the instrument.

V. DEMOGRAPHIC SECTION

A. The demographic information should be limited to only those variables that will be used to answer specific questions.
B. The demographic section should reveal the relationship of the respondent to the object of measurement.
C. Certain items in the demographic sections should be optional.
D. Language in the demographic section should be appropriate to the language level of the respondents.
E. Demographic items should be grouped together.

VI. WRITING ITEMS

A. Rating scale
   1. The stems of rating scale items should be written unidimensionally.
   2. The response sets to rating scales should be written unidimensionally.
   3. The response set of a rating scale should be tied logically to the stem.
   4. The level of specificity of the stem should answer the user's needs.
   5. Use the rating scale item only when a more direct method is not available.
   6. Directions for using the rating scale should be included, along with appropriate examples.
      a. Directions for a numerical rating scale
      b. Directions for a graphic rating scale
      c. Directions for a Likert scale
   7. The language used in rating scale items and responses should be written at the respondents' level of understanding.
   8. The type of rating scale format selected should be easily understood.
   9. Rating scale items should not elicit biased responses.
   10. The response sets for a rating scale should present choices along similar degree continuums.
   11. Avoid using global terms in response sets.
   12. Rating scales should include from three to seven categories.
   13. Choose the format for your rating scales that corresponds to the sophistication of respondents.
   14. Allow for a "not applicable" response when appropriate.
   15. Choose an analysis strategy appropriate to the data produced and the uses for which the questionnaire was designed.

B. Qualitative selection items
   1. State clearly and unambiguously the stem and responses.
   2. Use language appropriate to the respondents.
   3. State the stem unidimensionally.
   4. Write an exhaustive response set.
   5. The response categories should be mutually exclusive.
   6. Supply directions for difficult items.

C. Supply items
   1. Whenever possible, appropriate units should be specified for supply items.
   2. Provide sufficient space for responses.

VII. STRUCTURE AND FORMAT

A. Items should be grouped according to item types or similarity of content.
B. Group items within sections according to ease with which they can be answered.
C. Length of the instrument should relate to respondents and purpose.
D. The instrument should be clearly reproduced.

REFERENCES

Aiken, L.R., Jr. "Procedures and Problems in Designing a College Course Evaluation and Questionnaire." College and University. 1975, 50:247-253.

Barron's Profiles of American Colleges, Vol. 1: Description of the Colleges, 9th ed. Woodbury, NY: College Division of Barron's Educational Series, Barron's Educational Series, Inc., 1974.

Berdie, D.R. "Questionnaire Length and Response Rate." Journal of Applied Psychology, 1973, 58:278-280.

Guilford, J.P. Psychometric Methods. New York: McGraw-Hill Book Company, 1954.

Henerson, M.E., Morris, L.L., and Fitz-Gibbon, C.T. How to Measure Attitudes. Beverly Hills: SAGE Publications, 1978.

Johnson, W.R., et al., "Effects of Alternative Positioning of Open-Ended Questions in Multiple-Choice Questionnaires." Journal of Applied Psychology. 1974, 59:219-221.

Kohlan, R.G. "Comparison of Faculty Evaluations Early and Late in the Course: Instructor Evaluation Questionnaire." Journal of Higher Education, 1973, 44:587-595.

Kornhauser, A. and Sheatsley, P. "Questionnaire Construction and Interview Procedure." In C. Selltiz, et. al., Research Methods in Social Relations. New York: Holt, Rinehart and Winston, 1959.

Matell, M.S. and Jacoby, J. "Is There an Optiomal Number of Alternatives for Likert Scale Items? Study 1: Reliability and Validity." Educational and Psychological Measurment. 1971, 31:657-674.

Musumeci, M., and Koen, S. Parent Questionnaire. Yorktown Heights, NY: Center for Resource Management, 1980.

Payne, S. L. The Art of Asking Questions. Princeton, NJ: Princeton University Press, 1954.

Remmers, H.H. "Rating Methods in Research on Teaching." In N.L. Gage (Ed.), Handbook of Research on Teaching. Skokie, IL: Rand McNally, 1963.

Whitney, D.R. and Feldt, L.S. "An[illegible]zing Questionnaire Results: Multiple Tests by Hypotheses and Multivariate Hypotheses." Educational and Psychological Measurement. 1973, 33:365-380.